

Caso práctico TEMA 4

Enrique y Ana se conocieron hace 3 años y han decidido irse a vivir juntos.

Les encantan los perros, así que además están pensando en adoptar uno. Enrique, que es programador, creó hace unos días un sencillo programa en Python para poder dar un vistazo rápido a las razas que más le gustan, a modo de DataFrame de Pandas en el que contempla tanto el nombre de la raza como una puntuación de 0 a 10 indicando cuánto le gusta esa raza.

Ana está de acuerdo en adoptar una de las razas que Enrique tiene contempladas, así que le ha dicho cuánto le gusta cada una de ellas (de 0 a 100) para que Enrique lo integre en su DataFrame. Así podrán calcular medias para decidir qué raza les gusta más a los dos.

¿Qué te pedimos que hagas?

A continuación tienes la primera parte del programa de Enrique. Accede al intérprete de Python en trinket.io, pega el código fuente que te entregamos y ejecútalo (con el botón del triángulo negro de la parte superior).

```
import pandas as pd
indice = ['uno', 'dos', 'tres', 'cuatro', 'cinco']
columnas = ['Raza', 'Puntos']
lista = [['Caniche', 8.1], ['Bulldog', 7.3], ['ChowChow', 7.6],
['Chihuahua', 9.0], ['Labrador', 9.3]]
print('-----')
print('--- Apartado 1 ---')
print('-----')
```

Utiliza la información que encontrarás en los contenidos de la unidad para crear un documento en el que escribirás el código fuente que Enrique ha necesitado seguir escribiendo para llegar a las soluciones correspondientes a lo que se te dice en los siguientes apartados.

Al final de los apartados te damos toda la salida por pantalla que deberá tener el programa, para que sepas con claridad si estás consiguiendo lo que se te solicita. Tendrás que usar constantemente la función *print* para ello.

Apartado 1:

Crea un DataFrame llamado *df* a partir de la lista de razas y puntuaciones otorgadas por Enrique, usando el índice, las columnas que se te proporcionan.

Apartado 2:

Sistemas de Big Data

Ana le entrega sus puntuaciones a Enrique y éste las transcribe en forma de la siguiente lista:

```
puntos_ana = [61,75,82,95,99]
```

Añade la columna con etiqueta 'PuntosAna' al DataFrame.

Apartado 3:

Como las puntuaciones de Ana han entrado en valores de 0 a 100 y las de Enrique estaban de 0 a 10, divide la columna 'PuntosAna' entre 10 de modo que todas las puntuaciones queden en la misma escala.

Apartado 4:

Enrique y Ana han estado viendo una revista de animales y han decidido añadir otras dos razas más, por lo que han creado la siguiente lista ya con sus puntuaciones integradas.

```
lista2 = [['Samoyedo',9.2,8.9],['Pinscher',8.1,6.7]]
```

Añade esas dos nuevas filas al DataFrame. Fíjate en que se han quedado sin ideas para los índices y han decidido usar 'nuevo1' y 'nuevo2' para las nuevas filas.

Apartado 5:

Crea la columna 'Media' para poder ver la puntuación media para cada raza.

Apartado 6:

Enrique y Ana deciden que no tendrán un bulldog. Elimina la fila cuyo índice es 'dos'.

Apartado 7:

Viendo las puntuaciones medias, Enrique y Ana deciden que la decisión va a estar entre chihuahua, labrador y samoyedo, por lo que deciden obtener sólo esas filas del DataFrame (en concreto desde la posición 2 hasta la 5 -no incluida-).

Apartado 8:

Por último, sólo por curiosidad, deciden ver información estadística sobre las filas que les han quedado.

A continuación tienes la salida esperada del programa:

```
-----  
--- Apartado 1 ---  
-----
```

Sistemas de Big Data

```

Raza      Puntos
uno       Caniche      8.1
dos       Bulldog      7.3
tres      ChowChow     7.6
cuatro    Chihuahua    9.0
cinco     Labrador     9.3

-----
--- Apartado 2 ---
-----
Raza      Puntos      PuntosAna
uno       Caniche      8.1          61
dos       Bulldog      7.3          75
tres      ChowChow     7.6          82
cuatro    Chihuahua    9.0          95
cinco     Labrador     9.3          99

-----
--- Apartado 3 ---
-----
Raza      Puntos      PuntosAna
uno       Caniche      8.1          6.1
dos       Bulldog      7.3          7.5
tres      ChowChow     7.6          8.2
cuatro    Chihuahua    9.0          9.5
cinco     Labrador     9.3          9.9

-----
--- Apartado 4 ---
-----
Raza      Puntos      PuntosAna
uno       Caniche      8.1          6.1
dos       Bulldog      7.3          7.5
tres      ChowChow     7.6          8.2
cuatro    Chihuahua    9.0          9.5
cinco     Labrador     9.3          9.9
nuevo1    Samoyedo     9.2          8.9
nuevo2    Pinscher      8.1          6.7

-----
--- Apartado 5 ---
-----
Raza      Puntos      PuntosAna      Media
uno       Caniche      8.1          6.1          7.10
dos       Bulldog      7.3          7.5          7.40
tres      ChowChow     7.6          8.2          7.90
cuatro    Chihuahua    9.0          9.5          9.25
cinco     Labrador     9.3          9.9          9.60
```

Sistemas de Big Data

```
nuevo1    Samoyedo    9.2    8.9    9.05
nuevo2    Pinscher    8.1    6.7    7.40
-----
--- Apartado 6 ---
-----
      Raza      Puntos    PuntosAna    Media
uno      Caniche    8.1      6.1      7.10
tres     ChowChow    7.6      8.2      7.90
cuatro   Chihuahua    9.0      9.5      9.25
cinco    Labrador    9.3      9.9      9.60
nuevo1    Samoyedo    9.2      8.9      9.05
nuevo2    Pinscher    8.1      6.7      7.40
-----
--- Apartado 7 ---
-----
      Raza      Puntos    PuntosAna    Media
cuatro   Chihuahua    9.0      9.5      9.25
cinco    Labrador    9.3      9.9      9.60
nuevo1    Samoyedo    9.2      8.9      9.05
-----
--- Apartado 8 ---
-----
      Puntos    PuntosAna    Media
count    3.000000    3.000000    3.000000
mean     9.166667    9.433333    9.300000
std       0.152753    0.503322    0.278388
min       9.000000    8.900000    9.050000
25%      9.100000    9.200000    9.150000
50%      9.200000    9.500000    9.250000
75%      9.250000    9.700000    9.425000
max       9.300000    9.900000    9.600000
```